

1. We consider the problem, of solving

$$Ax = b$$

where we require  $A^{-1}$ . This computation is intensive, so we wish to approximate  $A^{-1}$  by solving the problem,

$$By = x$$

Where  $B$  is our goal. As a measure of error, we consider the MSE formula,

$$MSE = \sum (y - \hat{y})^2 / N$$

Here,

$$MSE = \sum (By - x)^2 / N$$

Which may be computed as the L2 norm of the vector,  $\langle By - x \rangle$ .